EVALUATION OF KIEFER'S VARIANCE BOUND IN A NON REGULAR CASE

BY

THOMAS POLFELDT

TECHNICAL REPORT NO. 65
AUGUST 8, 1969

PREPARED UNDER CONTRACT Nonr-225(72)
(NR-042-993)

FOR
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
EVALUATION OF KIEFER'S VARIANCE BOUND IN A
NON REGULAR CASE

by

Thomas Polfeldt

TECHNICAL REPORT NO. 65
August 8, 1969

PREPARED UNDER CONTRACT Nonr-225(72)
(NR-042-993) FOR
OFFICE OF NAVAL RESEARCH

This research was carried out with the support
of the University of Lund and the Sweden-American
Foundation.

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
EVALUATION OF KIEFER'S VARIANCE BOUND IN A
NON REGULAR CASE

by

Thomas Polfeldt

1. Introduction and Summary.

We consider the following situation. In a one-sided distribution
with c.d.f. \( F(x-\theta) \),

\[
F(x-\theta) = 0 \quad (x \leq \theta) \quad F(x-\theta) > 0 \quad (x > \theta)
\]

the location parameter \( \theta \) is estimated by means of \( n \) independent
observations. The density \( f(x-\theta) = f(y) \) exists and varies regularly
(at \( y = 0 \)) with exponent \( c-1 \):

\[
\lim_{y \to 0} f(ky)/f(y) = k^{c-1} \quad (\text{all } k > 0).
\]

The Kiefer (1952) variance bound for unbiased estimates \( t \) of \( \theta \) is,
in this situation,

\[
(1) \quad \inf_{t} V_{\theta}(t) = \sup_{\lambda} \frac{\left( \int_{0}^{\infty} h \, d\lambda(h) \right)^{2}}{\int_{(n)} \left( \int_{1}^{\infty} f(y_{1} - h) \, d\lambda(h) \right)^{2} \left( \prod_{1}^{n} f(y_{1}) \right)^{-1} \prod dy_{1}}
\]

where \( \lambda \) is the difference of two probability distribution functions.
(The more general case of \( \lambda \) as any function of bounded variation can
be reduced to the previous - cf. Kiefer (1952).)

In an earlier paper (Polfeldt (1967)), the author has shown that
this bound equals zero when \( c < \frac{1}{2} \). In this paper, we shall find a
closed expression for the limit (as $n \to \infty$) of \( \inf \mathcal{V}(t)/[F^{-1}(1/n)]^2 \), when \( \frac{1}{2} < c < 2 \) (section 2; see (4)). In section 3, we give an account of a numerical evaluation of this bound. The case $c = 2$ is treated in section 4. When $c > 2$, the situation can be considered as regular - cf. Blischke et al. (1968) and Blischke (1969b) - and the theory of Chernoff, Castwirth and Johns (1967) can be modified so as to apply. (No proof of this is given here, however.) A discussion of the results, with particular regard to the attainability of Kiefer's bound (cf. Barankin (1949), M. M. Rao (1965)) and to the author's previous work on best linear combinations of order statistics (Polfeldt (1969b)), is found in section 5.

An extension of the results to left one-sided distributions ($F(x-\theta) < 1$ (x < \theta), $F(x-\theta) = 1$ (x \geq \theta)) is obvious. An extension to biased estimates, or mean square error, is straightforward; some conditions on the bias will be necessary.

A slowly varying function (at $y = 0$) is a continuous, nonnegative function $L$ that satisfies

\[ \lim_{y \downarrow 0} L(ky)/L(y) = 1 \text{ (all fixed } k > 0) \]

Karamata (1930) (cf. Feller (1966), VIII.9) showed that $L(y)$ varies slowly at $y = 0$ if and only if

\[ L(y) = a(y) \exp \int_y^1 t^{-1} \epsilon(t) dt \]  

where $a(y) \to A \neq 0$ and $\epsilon(y) \to 0$ as $y \to 0$.

If $a(y) \equiv A$, $L(y)$ is a normalized slowly varying function.

**Definition.** We shall call $L(y)$ weakly normalized if there exists a
monotone function \( \Lambda(h) \) such that as \( h \to 0, \Lambda(h) \to \infty, h\Lambda(h) \to 0, \) and

\[
(3) \quad a(y(h)-h)/a(y(h)) = 1+o(h/y(h)) \quad (h \to 0)
\]

for any \( y(h) \) with \( h < y(h) < h\Lambda(h) \).

If \( f(y) = y^pL(y) \), \( f(y) \) varies regularly at \( y = 0 \) with exponent \( p \).

2. The Variance Bound.

**Theorem 1.** If

(i) there is a fixed number \( h_0 \) such that

\[
H = \{h|0 < h < h_0\} \cap \{h|h \neq 0, f(y) = 0 \implies f(y-h) = 0\} \cap
\]

\[
\{h|h+h \text{ is a possible parameter value}\}
\]

(ii) \( f(y) = cy^{c-1}L(y) \) varies regularly with exponent \( c-1 \), \( (c > 0) \) and the slowly varying function \( L(y) \) is weakly normalized

(iii) \( f(y-h)/f(y) = 1+O(h/y) \) (all \( y, h\Lambda(h) < y < \eta \), all \( h \in H \), \( \eta > 0 \) fixed, \( \Lambda(h) > 1 \) as in (3))

(iv) \( 0 < \lim_{h \to 0} h^{-2} \int_{\eta}^{\infty} \left(f(y-h)/f(y)-1\right)^2 f(y) dy \)

then, for \( 1/2 < c < 2 \), the Kiefer variance bound (1) fulfills

(as \( n \to \infty \))

\[
(4) \quad \inf V(t) \geq \mu_0 = a_n^2(1+o(1)) \sup_{\lambda} \frac{(\int_{0}^{\infty} s \, d\lambda(s))^2}{\int_{0}^{\infty} \int_{0}^{\infty} A(s,t) d\lambda(s) d\lambda(t)}
\]

3
where
\[ A(s,t) = \exp[\max(s^c, t^0) J(\min(s, t)/\max(s, t))]. \]

and

(5) \[ J(x) = x^{\Gamma(1+c)} \Gamma(2-c) \mathbf{F}_1(1-c, 2-c; 2; x). \]

For the proof, we need the following

**Lemma 1.** If \( L(y) \) is a weakly normalized slowly varying function at \( y = 0 \), if \( \Lambda(h) \) is defined as in (3), and if condition (iii) of the theorem holds, then for all \( y, h < y < \eta \)

\[ (1 - \frac{h}{y})^\kappa \leq \frac{L(y-h)}{L(y)} \leq (1 - \frac{h}{y})^{-\kappa}, \]

if \( h < y < h\Lambda(h) \), then \( \kappa = o(1) \) \( (h \to 0) \).

**Proof of Lemma 1.** We use the representation (2). Then

\[ \frac{L(y-h)}{L(y)} = \frac{a(y-h)}{a(y)} \exp \int_{y-h}^y t^{-1} \varepsilon(t) dt. \]

By the definition (3), if \( h < y < h\Lambda(h) \), \( a(y-h)/a(y) = 1 + o(h/y)^{\xi(1+h/y)^{\pm \kappa'}} \), where \( \kappa' \to 0 \) as \( h \to 0 \). On the other hand, when \( h\Lambda(h) < y < \eta \), condition (iii) gives \( a(y-h)/a(y) = 1 + o(h/y)^{\xi(1+h/y)^{\pm \kappa'}} \) with some \( \kappa' \).

Furthermore, for the integral in the exponent we have

\[ \int_{y-h}^y \leq \int_{y-h}^y \sup_{y-h < t < y} |\varepsilon(t)| \log y/(y-h) \leq \kappa'' \log y/(y-h) \]

(6) \[ \int_{y-h}^y \geq \inf_{y-h < t < y} \varepsilon(t) \log y/(y-h) \geq -\kappa'' \log y/(y-h), \]
if \( k'' = \max(\sup |e(t)|, |\inf e(t)|) \). When \( h < y < h \Lambda(h) \), the greatest possible argument for \( e \) tends to zero as \( h \to 0 \), and therefore \( k'' \to 0 \), or \( k'' = o(1) \). When \( h \Lambda(h) < y < \eta \), we can only assert \( k'' = o(1) \).

Raising \( e \) to the powers in (6) now gives the lemma.

**Remark.** The bounds

\[(7) \quad K_{\xi} (1-h/y)^{\xi} \leq \mathbb{L}(y-h)/\mathbb{L}(y) \leq K_{\xi}^{-1} (1-h/y)^{-\xi}\]

with some \( K_\xi \), with \( K_\xi = 1 \) if \( h < y < h \Lambda(h) \), and with \( \xi = \xi(\eta) \) arbitrarily small, can be obtained in a similar manner.

**Proof of Theorem 1.** We restrict our attention to the case \( 1/2 < c < 2 \).

Changing the integration order in the denominator of Kiefer's bound, we obtain

\[(8) \quad \int_0^\infty \int_0^\infty [\phi(h,k)]^n d\lambda(h) d\lambda(k)\]

with

\[\phi(h,k) = \int_{\max(h,k)}^{\infty} f(y-h)f(y-k)/f(y) dy .\]

Writing \( w = \max(h,k) \), \( \nu = \min(h,k) \), and observing that \( f(y-h)f(y-k) = f(y-w)f(y-v) = (f(y-w)-f(y))(f(y-v)-f(y)) + f(y)(f(y-w)+f(y-v)) - f^2(y) \),

we obtain

\[(9) \quad \phi(h,k) = 1 + F(w) - F(w-v) + \int_w^\infty \left( \frac{f(y-w)}{f(y)} - 1 \right) \left( \frac{f(y-v)}{f(y)} - 1 \right) f(y) dy .\]

Now, for the integral over \( \eta < y < \infty \) we have, according to condition (iv),
2\left[\int_{\eta}^{\infty} - \int_{\eta}^{\infty} \left(\frac{f(y-w)}{f(y)} - 1\right)^2 f(y) dy + \int_{\eta}^{\infty} \left(\frac{f(y-v)}{f(y)} - 1\right)^2 f(y) dy\right]
\begin{align*}
\Rightarrow \int_{\eta}^{\infty} \left(\frac{f(y-w)}{f(y)} - 1\right)^2 f(y) dy + \int_{\eta}^{\infty} \left(\frac{f(y-v)}{f(y)} - 1\right)^2 f(y) dy.
\end{align*}

(10)
\begin{align*}
= O(w^2 + v^2) \quad (w, v) \to 0.
\end{align*}

For small arguments \( y \), condition (ii) implies

(11)
\begin{align*}
F(y) = y^c L(y)(1 + o(1)) \quad (y \to 0).
\end{align*}

Introducing \( f(y) = cy^{c-1} L(y) \) in the integral over \( w < y < \eta \) in \( \Phi(h, k) \), (9), we obtain

(12)
\begin{align*}
\int_{\eta}^{\infty} \left(\frac{1-y}{y} - 1\right)^2 \frac{L(y-w)}{L(y)} - 1\right)^2 \frac{L(y-v)}{L(y)} - 1\right) cy^{c-1} L(y) \frac{L(w)}{L(w)} dy.
\end{align*}

Here, we need \( c > 1/2 \) for the integral to converge at \( y = w \).

We divide the integration interval into \( (w, w \Lambda(w)) \) and \( (w \Lambda(w), \eta) \)
(where \( \Lambda \to \infty \), \( w \Lambda \to 0 \) as \( w \to 0 \)) and apply lemma 1, including the remark (7). We also change variables, \( w/y = s \), and write \( v/w = x \), to obtain

\begin{align*}
\int_{\eta}^{\infty} \left[ (1-s)^{c-1-k} - 1\right][(1-sx)^{c-1+k} - 1]cs^{-1+k} dswL(w).
\end{align*}

The last of these two integrals is (since \( c < 2 \))

\begin{align*}
0(s^{1/\Lambda(w)} s^2 x^{c-1+k} c s^{-1+k} ds \cdot w^c L(w))
\end{align*}

\begin{align*}
= 0(\ln^2 w \Lambda(w)^{-2+c+k}) = o(F(w)),
\end{align*}

while in the first integral, \( \kappa \to 0 \) as \( w \to 0 \), and \( K_5 = 1 \), and therefore,
\[(13) \quad \int_{1/\Lambda(w)}^{1} \frac{1}{(1-s)^{c-1}-1} \frac{1}{((1-sx)^{c-1}-1)cs^{c-1}ds} F(w)(1+o(1)) = (w \to 0).\]

It is easy to show (e.g. by a series expansion of the factor \((1-sx)^{c-1}-1\)) that the right member of (13) equals (cf. (5))

\[(14) \quad F(w)(1+o(1))[(1-x)^{c-1}+x \frac{1}{\Gamma(1-c) \Gamma(2-c)} \frac{1}{\alpha^c} F_1(1-c,2-c;2;x)] = F(w)(1+o(1))[(1-x)^{c-1}+J(x)].\]

If we assemble the various parts of \(\Phi(h,k)\) from (9), (10) and (14), we find that

\[\Phi(h,k) = 1+F(w)J(x) + o(F(w)) + o(w^2 + v^2) + F(w)(1+o(1))[(1-x)^{c-1}] + F(w)-F(w-v).\]

The last line is simply

\[(1-x)^{c}F(w)(1+o(1)) - F(w-v).\]

Using (11) and the fact that \(x = v/w\), we obtain for small values of \(w\),

\[(1-x)^{c}w^{c}L(w)(1+o(1)) - w^{c}(1-x)^{c}L(w-v).\]

But lemma 1 can be applied to \(L(w-v)/L(w)\), to give

\[(15) \quad (1-x)^{c}w^{c}L(w)[1+o(1)] - (1-x)^{c}K.\]

When \(w \to 0\), we have \(K \to 0\), and so (15) is \(o(F(w))\). (For \(1-k<x<1\), we rely on the first factor, \((1-x)^{c}\).) Noting that when \(c<2\), also
0(u^2 + v^2) = o(F(w)) \text{, we obtain}

\[ \Phi(h, k) = 1 + F(w)J(x) + o(F(w)) \quad (w \to 0). \]

Now, introduce this expression in (8), and change the variables into

\[ h = a_n s, \quad k = a_n t; \quad \text{for convenience, write} \quad w = a_n W, \quad v = a_n V, \quad x = v/w = V/w. \]

Let \( W_o(n) \) depend on \( n \) in such a manner that \( W_o(n) \to \infty \ (n \to \infty), \)

\[ n^{-1}(W_o(n))^{2c+\delta} \to 0 \quad (n \to \infty) \quad (\text{where any} \ \delta > 0 \ \text{will do}). \]

Applying lemma, we find that for \( 0 < W < W_o(n), \)

\[ L(a_n W)/L(a_n) \geq W^{\pm K}, \]

where \( K \to 0 \) as \( n \to \infty. \) For these values of \( W, \)

\[ F(w) = F(a_n W) = a_n^c W^{\pm K} (L(a_n)/L(a_n)) \cdot L(a_n) = \]

\[ = W^{c+K}F(a_n)W^{-\pm K} = W^{c-K}/n, \]

and so

\[ (1 + F(w)J(x))^n = \exp[W^{c+K}J(x)](1 + o(W_o(n))^{2c+2K}/n)). \]

As long as the integrals converge, we then get (if \( \lambda(s) = \lambda(a_n s) \))

\[ \int_0^\infty \int_0^\infty [\Phi(h, k)]^n d\lambda(h) d\lambda(k) = \]

\[ = \int_0^{W_o} \int \exp[W^{c+K}J(x)]d\lambda(s)d\lambda(t)(1 + o(W_o^{2c+2K}/n)) \]

\[ + \int_{s > W_o} \int_{t > W_o} \Phi^n d\lambda(s)d\lambda(t) \]

\[ = \int_0^\infty \int \exp[W^{c+K}J(x)]d\lambda(s)d\lambda(t)(1 + o(1)) \quad (n \to \infty). \]
Since $\kappa$ can be arbitrarily small, the value of this integral will be

$$\int_0^\infty \exp(w^c J(x))d\lambda(s)d\lambda(t) (1+o(1)) \ (n \to \infty),$$

and the theorem is proved.

3. **Numerical Results.**

For the numerical solution of the supremum problem (4), we use the Rayleigh-Ritz procedure (cf. any textbook on numerical solution of variational problems, e.g. Berezin and Zhidkov (1965)). We also use the following

**Lemma 2.** The function $\lambda(s)$ that maximizes (4) can be discontinuous only at $s = 0$.

The proof will be found at the end of this section. In essence, $\lambda(s)$ is approximated by $\sum_0^P a_i G_i(s)$, where the $G_i$'s are chosen so that a) for any $\epsilon > 0$, and for $P$ large enough, there exist coefficients $[a_i]_0^P$ that make $|\lambda(s) - \sum_0^P a_i G_i(s)| < \epsilon$ uniformly in $s$, for the entire class of $\lambda$'s we want to consider - i.e., all continuous functions that are the difference of two probability measures, with a possible discontinuity at $s = 0$; and b) it is possible to calculate the quantities

$$\tag{16} b_{ij} = \int \int A(s,t)dG_i(s)dG_j(t).$$

We have chosen $G_0(s) = 0 \ (s < 0)$, $G_0(s) = 1 \ (s \geq 0)$, and $G_1(s) = \exp(-K_1 s^c) \ (s > 0)$, where the $K_1$ are positive numbers.
which will be stated later. In the integrals (16), we then have (for
\(i,j \geq 1\))
\[
\iint_{s < t} = \int_0^{\infty} \int_0^{1} \exp(t^c J(x))K_iK_j e^{2c-1}x^{c-1} \exp(-t^2(K_i^c + K_j^c))dxdt
\]
\[
= K_iK_j \int_0^1 [K_i^c + K_j^c - J(y^c)]^{-2}dy = I_{ij}
\]
(17)
provided that \(K_i^c + K_j^c - J(y^c) > 0 \) (0 < y < 1), or \(\min(K_i^c) > J(y^c)/(1+y)\).
Since this function of \(y\) increases in \((0,1)\) for all \(c < 2\), this condition amounts to
\[
\min(K_i^c) > J(1)/2 = r(2-c)r(2c-1)/(2r(c)).
\]
The \(K_i^c\) 's that have been used are in fact \(K_{i+4} = (1+i)J(1)/2\)
(i=1,2,\ldots) - these \(G_i^c\) 's themselves together with \(G_0^c(s)\) fulfill
the condition a) above - plus a few extra \(K\)-values: \(K_i = (1+2^{-5+i})J(1)/2\)
(i=1,2,3,4).

We thus have (from (16) and (17))
\[
b_{0i} = b_{io} = 1 \quad (i=0,1,\ldots)
\]
\[
b_{ij} = I_{ij} + I_{ji} \quad (i,j \geq 1).
\]

The integrals (17) have been calculated numerically. When we introduce
\(\sum_0^P a_iG_i(s)\) for \(\lambda(s)\) in (4), we have reduced the problem to finding (in
a matrix notation)
\[
\mu_o(p) = \sup_a t^D a/a^T b
\]
if \(a^T = (a_0,a_1,\ldots,a_p)\); \(D = (d_{ij}) = (\delta_i \cdot \delta_j)\) with \(\delta_o = 0,\)
\(d_i = \int_0^\infty sG_i(s) = r(1+\alpha)K_i^\alpha\) (i \geq 1), and \(B = (b_{ij})\).
This is a standard problem - $\mu_0^{(p)}$ is the largest number $\mu$ that satisfies

$$Da = \mu Ba,$$

and $a$ is the corresponding eigenvector of $B^{-1}D$. In our case, both $D$ and $B$ are symmetric and $B$ is positive definite, and there exist good techniques for solving this so-called generalized eigenvalue problem without considering explicitly the matrix $B^{-1}D$. In the computer program, we used the subroutines GEIGEN and EIGEN from the collection of scientific subroutines distributed by the Argonne National Laboratories. The theoretical background of these subroutines is described, i.e., in Wilkinson (1965). (pp. 337ff., 290ff., 299ff.)

The following values for $\mu_0$ were calculated (Table 1). Generally, less than 10 G's were needed to obtain convergence of $\mu_0^{(p)}$ to $\mu_0$. As a comparison, we quote the number $V(\infty) = (2\alpha - 1)\Gamma(1/2\alpha)/2\alpha (\alpha = 1/c)$ which represents the best variance attainable by an estimator which is linear in the (ordered) observations - cf. Polfeldt (1969b).
TABLE 1. Values of Kiefer's Variance Bound

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.51</td>
<td>0.53</td>
<td>0.55</td>
<td>0.60</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>c</td>
<td>1.96</td>
<td>1.89</td>
<td>1.82</td>
<td>1.67</td>
<td>1.43</td>
<td>1.33</td>
</tr>
<tr>
<td>μ₀</td>
<td>0.0180</td>
<td>0.0496</td>
<td>0.0788</td>
<td>0.1494</td>
<td>0.3083</td>
<td>0.3992</td>
</tr>
<tr>
<td>V(∞)</td>
<td>0.0198</td>
<td>0.0581</td>
<td>0.0951</td>
<td>0.1836</td>
<td>0.3549</td>
<td>0.4431</td>
</tr>
<tr>
<td>μ₀/V(∞)</td>
<td>0.9091</td>
<td>0.8539</td>
<td>0.8284</td>
<td>0.8135</td>
<td>0.8688</td>
<td>0.9009</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.90</td>
<td>1.00</td>
<td>1.10</td>
<td>1.20</td>
<td>1.25</td>
<td>1.30</td>
</tr>
<tr>
<td>c</td>
<td>1.11</td>
<td>1.00</td>
<td>0.91</td>
<td>0.83</td>
<td>0.80</td>
<td>0.77</td>
</tr>
<tr>
<td>μ₀</td>
<td>0.7307</td>
<td>1.0000</td>
<td>1.2957</td>
<td>1.5987</td>
<td>1.7439</td>
<td>1.8790</td>
</tr>
<tr>
<td>V(∞)</td>
<td>0.7451</td>
<td>1.0000</td>
<td>1.3222</td>
<td>1.7390</td>
<td>1.9940</td>
<td>2.2874</td>
</tr>
<tr>
<td>μ₀/V(∞)</td>
<td>0.9807</td>
<td>1.0000</td>
<td>0.9800</td>
<td>0.9193</td>
<td>0.8746</td>
<td>0.8215</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1.40</td>
<td>1.45</td>
<td>1.50</td>
<td>1.55</td>
<td>1.00</td>
<td>1.65</td>
</tr>
<tr>
<td>c</td>
<td>0.71</td>
<td>0.69</td>
<td>0.67</td>
<td>0.65</td>
<td>0.63</td>
<td>0.61</td>
</tr>
<tr>
<td>μ₀</td>
<td>2.0986</td>
<td>2.1724</td>
<td>2.2161</td>
<td>2.2243</td>
<td>2.1936</td>
<td>2.1313</td>
</tr>
<tr>
<td>V(∞)</td>
<td>3.0177</td>
<td>3.4720</td>
<td>4.0000</td>
<td>4.6150</td>
<td>5.3327</td>
<td>6.1719</td>
</tr>
<tr>
<td>μ₀/V(∞)</td>
<td>0.6954</td>
<td>0.6257</td>
<td>0.5540</td>
<td>0.4820</td>
<td>0.4113</td>
<td>0.3454</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>1.75</td>
<td>1.80</td>
<td>1.85</td>
<td>1.90</td>
<td>1.95</td>
<td>1.97</td>
</tr>
<tr>
<td>c</td>
<td>0.57</td>
<td>0.56</td>
<td>0.54</td>
<td>0.53</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>μ₀</td>
<td>1.8691</td>
<td>1.4730</td>
<td>0.9044</td>
<td>0.2115</td>
<td>0.0344</td>
<td>0.0158</td>
</tr>
<tr>
<td>μ₀/V(∞)</td>
<td>0.2250</td>
<td>0.1524</td>
<td>0.0803</td>
<td>0.0161</td>
<td>0.0022</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

(The diagrams on page 13 were drawn on the basis of these data.)
For some c-values, the coefficients \( \{a_i\} \) of the approximation \( \sum_o^D a_i G_i(s) \) of the maximizing function \( \lambda \) are listed in Table 2. The coefficients are normed so that

\[
\sum (a_i > 0) a_i = 1, \quad \sum (a_i < 0) a_i = -1,
\]

thus offering one possibility of expressing \( \lambda \) as a difference of two probability measures. For some values of \( c \) (close to \( c = 1 \)), it was noted that \( \sum_o^D a_i G_i(s) \) was very well approximated by \( (1-\exp(-Kc^s)) - G_0(s) \) (if the norming \( a_o = -1 \) is carried out). In the program, it was found convenient to use the formula \( K_i = (1+2^{-5+1})J(1)/2 \) throughout, starting with \( i = 8 \). Thus, some of the \( G_i(s) \) were not actually included in the calculations reported here. (In others they were, with no change in the results of Table 1.)
**TABLE 2**

$a_i$ is the coefficient for $G_i(s) = 1 - \exp(-K_i s^C)$; $K_i$ as on page 10; $J=J(1)$

<table>
<thead>
<tr>
<th></th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
<th>$a_8$</th>
<th>$a_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.379</td>
<td>0.251</td>
<td>0.000</td>
<td>0.755</td>
<td>0.771</td>
<td>0.781</td>
<td>0.205</td>
<td>0.178</td>
<td>0.214</td>
<td>0.228</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.000</td>
<td>0.801</td>
<td>0.385</td>
<td>0.636</td>
<td>0.796</td>
<td>0.417</td>
<td>0.154</td>
<td>0.076</td>
<td>0.213</td>
<td>0.039</td>
</tr>
</tbody>
</table>

15
We conclude this section by the

**Proof of Lemma 2.** From the theory of the calculus of variations, it is known that the maximizing \( \lambda \) in (4) must fulfill

\[
\int_0^\infty A(s,t)d\lambda(s) = \text{const} \cdot t \quad (\text{for all } t).
\]

Now assume that \( \lambda = \lambda_c + \lambda_d \), where \( \lambda_c \) is absolutely continuous with derivative \( \ell_c \), and \( \lambda_d \) is discrete, with discontinuities of size \( S_i \) at the points \( s = \sigma_i \) \((i=1,2,\ldots)\). Then, we can write (18) as

\[
\text{const} \cdot t = \int_0^\infty A(s,t)\ell_c(s)ds + \sum_i S_i A(\sigma_i,t).
\]

Let \( D \) denote differentiation with respect to \( t \). Then,

\[
D(\text{const} \cdot t) = \text{const} = \int_0^\infty D A(s,t)\ell_c(s)ds + \sum_i DA(\sigma_i,t).
\]

Now,

\[
DA(s,t) = \begin{cases} 
 s^{c-1}J(t/s)\exp(s^c J(t/s)) & 0 < t < s \\
 s^{c-1}(sJ(s/t)-s^{c-1}J'(s/t))\exp(t^c J(s/t)) & t > s > 0.
\end{cases}
\]

It is easy to establish that \( J(x) \sim J(1)-K(1-x)^{2c-1} \) \((x \to 1)\), and that \( J'(x) \sim K(1-x)^{2c-2} \) \((x \to 1)\). For \( c < 1 \), the right side of (19) will then be the sum of a continuous function \((J'(t/s))\) and \(J'(s/t)\) are integrable at \( s=t \) and a discontinuous function

---

16
(at \( t = \sigma_i \)), and (19) cannot be satisfied at \( t = \sigma_i \), unless \( \delta_i = 0 \). Only for \( \sigma_i = 0 \) does this argument not apply, since \( DA(0, t) = D(1) = 0 \). When \( c > 1 \), we can perform one more differentiation with respect to \( t \). In this case also \( J''(x) \) and \((J'(x))^2\) are integrable at \( x = 1 \), and the result follows in the same way as before.

4. The Case \( c = 2 \).

For easier display, we shall restrict our attention to \( f(y) = y L_1(y) \) with a normalized slowly varying \( L_1(y) \). A scrutiny of the proof will reveal that some (but not all) weakly normalized \( L_1(y) \) can be covered by the same proof.

The case where \( T(x) \to \text{const} (x \to 0) \) (cf (iii') below) can be considered regular in as much as the best variance is of the exact order \( n^{-1} \) in this case; much stronger results can be found along the lines of Blischke et al. (1968) and Blischke (1969b). The author intends to treat this problem in another context.

**Theorem 2.** Under the conditions (i) and (iv) of theorem 1, if also

(ii') \( f(y) = 2yL(y) = y L_1(y) \) with normalized slowly varying \( L_1(y) \)

(iii') \( T(x) = \int_x y^{-1} L_1(y) dy \to \infty \) (\( x \to 0 \))

and if \( G(x) = x^{2}T(x), \ A_n = G^{-1}(1/n) \),

then the Kiefer variance bound (1) is

\[
\inf V(t) = A_n^2(1 + o(1)) \sup_\lambda \frac{\left( \int_0^\infty s d\lambda(s) \right)^2}{\int_{\text{est}} d\lambda(s) d\lambda(t)} = A_n^2(1 + o(1))
\]
Proof. The proof of theorem 1 can be followed up to and including (12). The integration domain is divided into \((w, k(w))\) and \((k(w), \eta)\), where \(k\) fulfills

\[
\begin{align*}
    \text{(20)} & \quad k(w) > w, \quad k(w) \to 0 \text{ as } w \to 0, \\
    & \quad T(k(w))/T(w) \to 0 \text{ as } w \to 0.
\end{align*}
\]

We apply lemma 1 to \(L(y-w)/L(y)\) and \(L(y-v)/L(y)\) in (12). Retaining the meaning of \(\kappa''\) from (6), we obtain

\[
\int_0^1 \left(1 - \frac{y}{y} \right)^{1+k''} - 1 \left(1 - \frac{y}{y} \right)^{1+k''} - 1 \right) y L_1(y) \, dy.
\]

In \((w, k(w))\), \(k'' \to 0\) (since \(k(w) \to 0\)), and thus

\[
\begin{align*}
    \int_w^k \frac{k(w)}{w} \, dy & = \int_w^k \frac{k(w)}{w} \, dy = 1 + o(1) \quad (w \to 0) \\
    & = w v T(w) (1 + o(1))
\end{align*}
\]

(due to the conditions (20)). In \((k(w), \eta)\), on the other hand, we obtain for the same reason

\[
\int_0^k \frac{\eta}{k(w)} \, dy = o(\int_0^\eta \frac{y}{k(w)} \, dy) = o(v w T(w))
\]

Thus, from (9),

\[
\Phi(h, k) = 1 + o(w^2 L_1(w) + (w-v)^2 L_1(w-v)) + w v T(w) (1 + o(1)) + o(v^2 + w^2)
\]

\[
= 1 + w v T(w) + o(v w T(w)) \quad (w \to 0).
\]

18
Introducing \( x, V, W \) by \( v = xw \), \( w = A_n W \), \( v = A_n V \),

\[ \phi(h, k) = 1 + xW^2/n + o(xW^2/n) . \]

Introducing this in (8), which as we noted is valid also in this case, and proceeding as in the last part of the proof of theorem 1, we obtain the variance bound

\[ A_n^2(1 + o(1))(\int s d\bar{\lambda}(s))^2/\int e^{xW^2} d\bar{\lambda}(s) d\bar{\lambda}(t) . \]

But \( xW^2 = VW = \min(s, t) \cdot \max(s, t) = st \).

Finally, standard methods from the calculus of variation will give

\[ \sup_\lambda(\int_0^\infty s d\bar{\lambda}(s))^2/\int_0^\infty e^{st} d\bar{\lambda}(s) d\bar{\lambda}(t) = 1 \]

(with \( \bar{\lambda}(s) = \lim_{h \to 0} h^{-1}[G_o(s-h)-G_o(s)] \), where \( G_o(s) \) is defined as in section 2.) The theorem is proved.

For \( c = 2 \) (and \( T(x) \to \infty \)), the Kiefer bound is asymptotically attainable. The estimate that attains it is asymptotically normally distributed; it is derived by the author (Polfedt (1969b)).

For the calculation of \( A_n \), cf. Polfeldt (1969a).

5. Discussion of the Results.

A previous report by the author (Polfedt (1969b)) has already been mentioned, and its result on the variance \( V(\infty) \) of the best linear estimates is reported in Table 1. There can be two reasons for the discrepancy (cf. Figure 2) between the Kiefer bound, \( \mu_o \), and \( V(\infty) \).
One is that linear estimates are not the best in the non-regular situation we are dealing with, except in the special case \( c=1 \) and the border case \( c=2 \). This point of view is supported by the expression for the Pitman estimate,

\[
t_p = \int \prod f(x_i - \theta) d\theta / \prod f(x_i - \theta) d\theta ,
\]

which is generally not linear in the observations (or the ordered observations) — but it is known to be the best location invariant estimate. Blischke et al. (1968) have worked along these lines, but as far as the author can see, without yet coming below \( V(\omega) \). If the best location invariant estimate is non-linear, we cannot hope to approximate it by a linear estimate in the non-regular case, where loosely speaking the first few order statistics contain almost all the information. A further reason against the linear estimates is the fact first noticed by Blischke (1969a), that when \( c > 1 \), \( P(t > \min(x_i)) > K_1 > 0 \) for the best linear estimates \( t \).

Another possible reason for \( V(\omega) - \mu_o \) to be positive is that the Kiefer bound may not be attainable. It is, however, locally attainable. In fact, following Barankin (1949), though his proof does not apply directly, it is easy to show that for \( \theta = \theta_o \), the estimate

\[
t_L = \int \prod \left( f(x_i - \theta) / f(x_i - \theta_o) \right) d\lambda_o(\theta) + \theta_o
\]

(where \( \lambda_o \) is the function that gives supremum in (4)) has \( V_{\theta_o}(t_L) = \mu_o \).

The possibility of attaining the Kiefer bound globally must however be left open to doubt, on the following grounds. Firstly, we
require from our estimates $t$ that $E_\theta(t) = \theta$ — but, as is seen by going through the steps of Kiefer’s proof (Kiefer (1952)), when we obtain the expression $\inf_t V_{\theta_o}(t) \geq \mu_o$, we only require $E_\theta(t) = \theta$ for $\theta > \theta_o$. Theoretically, for $\theta < \theta_o$, $E_\theta(t)$ could be anything — and the bound might be attainable only in this wider class of estimates. Secondly, and more importantly, the Kiefer bound tends to zero as $c \downarrow 1/2$. However, there are other bounds, e.g. the one constructed by the author (Polfeldt (1967), (1968)) that are positive for $c \leq 1/2$. For continuity reasons they are also positive for $c = 1/2 + \epsilon$, and it is conceivable that they may be greater than the Kiefer bound in a considerable part of $1/2 < c < 2$.

Thus, both the possible explanations for the difference between $V(\omega)$ and $\mu_o$ seem plausible. The author plans further work on these lines.

Acknowledgment.

I wish to thank Professors Herman Chernoff, Herman Rubin and M.V. Johns, Jr. for stimulating discussions, which to some extent are reflected in section 5 of this paper.
REFERENCES


Evaluation of Kiefer's Variance Bound in a Non Regular Case

For a one-sided distribution defined by the density $f(x-\theta) = (x-\theta)^{c-1}L(x-\theta)$ ($x > \theta$, $f(x-\theta)=0$, $x < \theta$), where $L(x-\theta)$ varies slowly at $x=\theta$, we estimate the unknown location parameter $\theta$ by means of $n$ independent observations. The estimate is unbiased. When $1/2 < c < 2$, the variance bound of Kiefer (Ann. Math. Stat., 23, 627-629 (1952)) is found to be equal to (as $n \to \infty$)

$$\mu_n = a_n^2 \frac{1}{n} \sup_\lambda \left( \int_0^\infty s \lambda(s)ds \right)^2 \int_0^\infty \Lambda(s,t)ds \lambda(s)dt$$

where $a_n$ is defined by $\int_0^\infty f(y)dy = 1/n$; $\lambda(s)$ is the difference between two probability measures, and $\Lambda(s,t) = \exp(\max(s^c,t^c)J(\min(s,t)/\max(s,t)))$ with $J(x) = x^c(1+c)(2-c)2^{c-1}(1-c,2-2c;2|x)$. The supremum is then calculated numerically. The case $c=2$ is treated separately. The case $c>2$ can be considered as regular; when $0 < c < 1/2$, the Kiefer bound is zero. Let $V(\omega)$ be the variance of the best linear combination of the ordered observations. The fact that $\mu_n/V(\omega) < 1$ (except for $c=1$ and $c=2$) is discussed.
Non regular estimation
Minimum variance estimation

<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROLE</strong></td>
<td><strong>WT</strong></td>
<td><strong>ROLE</strong></td>
<td><strong>WT</strong></td>
</tr>
</tbody>
</table>

**INSTRUCTIONS**

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Masking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, c & d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

   (1) "Qualified requesters may obtain copies of this report from DDC."

   (2) "Foreign announcement and dissemination of this report by DDC is not authorized."

   (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through Ed.

   (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through Ed.

   (5) "All distribution of this report is controlled. Qualified DDC users shall request through Ed."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (O), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.